

# 基于深度学习的混凝土结构多病害检测

饶勇成<sup>1</sup>, 韩晓健<sup>1</sup>, 肖 飞<sup>2</sup>, 孙思其<sup>1</sup>

(1 南京工业大学土木工程学院, 南京 211800;

2 江苏建研建设工程质量安全鉴定有限公司, 南京 211800)

**[摘要]** 本文针对混凝土常见病害, 提出了一种改进 Unet 网络对混凝土多病害进行检测的方法, 建立了一个包含混凝土裂缝、泛碱、剥落、露筋、孔洞数据集; 再裁剪原始图像, 选取 10000 张小块图像进行随机翻转, 共产生 13200 张病害数据集用于研究; 最后, 将数据集对改进的 Unet 网络进行训练、验证和测试, 实现了像素级病害区域识别。为验证网络模型性能, 将模型与原 Unet、基于 VGG16 的 Unet、FCN 和基于 ResNet50 的 FCN 进行了比较, 结果表明所改进的网络模型检测效果优于其他模型。为测试模型鲁棒性, 利用滑动窗口算法对不同环境下的混凝土病害进行全局检测, 取得了较好的检测效果。

**[关键词]** 深度学习; 语义分割; 混凝土; 病害检测

中图分类号: TU17

文献标识码: A

文章编号: 1002-848X(2021)S2-1439-07

## Multiple disease detection of concrete structures based on deep learning

RAO Yongcheng<sup>1</sup>, HAN Xiaojian<sup>1</sup>, XIAO Fei<sup>2</sup>, SUN Siqi<sup>1</sup>

(1 College of Civil Engineering, Nanjing Tech University, Nanjing 211800, China;

2 Jiangsu Jian Yan Civil Engineering Quality Appraisal Co., Ltd., Nanjing 211800, China)

**Abstract:** Aiming at the common diseases of concrete, this paper proposes an improved Unet network to detect multiple diseases of concrete, and establishes a data set containing concrete cracks, efflorescence, spalling, exposed tendons, and holes; afterwards, the original images are cropped and 10000 pieces are selected. The block images are randomly flipped to produce a total of 13200 disease data sets for research; finally, the data sets are used to train, verify and test the improved Unet network to realize pixel-level disease area recognition. In order to verify the performance of the network model, the model was compared with the original Unet, VGG16-based Unet, FCN and ResNet50-based FCN. The results show that the improved network model detection effect is better than other models. In order to test the robustness of the model, the sliding window algorithm is used to globally detect concrete diseases in different environments, and good detection results have been achieved.

**Keywords:** deep learning; semantic segmentation; concrete; disease detection

## 0 引言

混凝土作为桥梁、大坝、房屋等结构的组成部分, 在维持结构正常使用中发挥着重要作用。随着国民经济的发展, 这些基础设施都在不断增长, 且随着使用年限的增加, 受外部荷载以及环境温湿度、雨雪侵蚀等影响, 结构耐久性问题也不断涌现。在混凝土表面容易出现病害现象。常见的混凝土病害有裂缝、剥落、孔洞、露筋、蜂窝、麻面等。其中裂缝是病害发展的最初形态, 在结构检测当中裂缝检测十分重要, 微小裂缝不会影响结构的安全性, 但当裂缝达到一定限值时, 外部环境中的腐蚀物质(如水蒸气等)进入结构, 造成钢筋锈蚀, 体积膨胀, 使得混凝土裂缝进一步扩展, 进而发展为剥落、露筋等其他病害, 演化过程如图 1 所示。

传统的人工检测方法需要检测人员去现场采

集图像, 效率低且检测精度依赖于检测人员的主观经验, 具有不确定性。因此, 基于机器视觉的缺陷检测技术应运而生, 目前, 缺陷检测技术在土木领域主要应用于混凝土裂缝检测, 同时对多种病害检测的研究较少, 使得病害检测不够全面。主要的裂缝检测技术可分为两种, 一是基于数字图像处理的裂缝检测技术<sup>[1-2]</sup>, 一是基于深度学习的裂缝检测技术。基于数字图像处理的检测技术主要针对某特定病害设计相关算法, 依赖人工调参, 往往特定算法只能检测特定病害类型, 当图像存在噪音或光照等问题时, 鲁棒性较差。基于深度学习的检测技术主要通过构建卷积神经网络模型来实现自动化检测, 相较于传统算法无需手动提

**作者简介:** 饶勇成, 硕士, Email: 201861226136@njtech.edu.cn;  
**通信作者:** 韩晓健, 副教授, 硕士生导师, Email: hanxj@foxmail.com。

取特征, 并且检测效果和性能有了极大的改进。

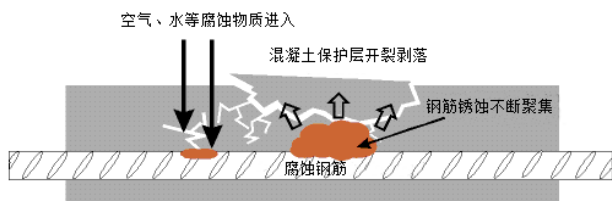


图1 混凝土剥落露筋演化过程

随着计算机技术的发展, 深度学习已广泛应用于计算机视觉领域, 产生了许多经典的检测算法, 如 FCN、SDD、Unet<sup>[3-5]</sup>等。研究人员将这一最新技术引入到土木工程领域来检测混凝土病害。韩晓健等<sup>[6]</sup>将深度学习和数字图像处理技术相结合, 提出了一种基于深度学习裂缝检测方法, 检测精度在验证集上平均正确率达到了 98.39%。Cha 等<sup>[7]</sup>利用深度卷积神经网络对识别混凝土裂缝, 采用滑动窗口技术检测任意大小图像, 实现了对裂缝分类。Park 等<sup>[8]</sup>使用两阶段深度学习方法先利用分割来提取路面, 随后通过分析提取路面内单位斑块来检测裂缝。李国耀等<sup>[9]</sup>提出一种结合形态学操作与对比度增强的预处理算法用于图像处理, 再利用基于裂缝几何特征算法实现了对裂缝区域的提取。虽然这些研究方法都能够准确的定位裂缝位置, 但无法对裂缝骨架进行提取, 不能直观观察裂缝形态。或者虽能够提取裂缝区域特征, 但应用的算法无法胜任不同环境下的检测, 鲁棒性不强。

基于深度学习的语义分割技术把图像中每个像素作为分类对象, 在像素级对图像中不同类别进行精准分割。2014 年 Long 等<sup>[3]</sup>提出的全卷积神经网络 (FCN) 是语义分割的开山之作, 之后 Ronneberge 等又提出 Unet<sup>[5]</sup>网络, 该网络可以端到端的训练网络模型, 从而可以输出与输入时间等尺度的图像。本文主要在前人研究裂缝基础上, 扩展了混凝土病害的种类, 针对 5 中常见的混凝土病害: 裂缝、剥落、孔洞、泛碱、露筋, 改进了 Unet 语义分割网络模型, 实现了对混凝土结构多种病害像素级识别检测。

## 1 网络结构

所提出的网络为 Res-Unet 网络。将 Unet 作为基本框架, ResNet-50 作为编码器。结构如图 2 所示。ResNet-50 编码器部分权重采用了在 ImageNet 上预训练的权重值, 通过迁移学习可以使得模型加速收敛。网络由一系列的功能层组成,

包括卷积层 (Conv)、批归一化层 (BN)、非线性激活层 (ReLU)、反卷积层 (Deconvolution)、全连接层 (Conv1x1) 等。

网络第一层为输入层, 输入维度为  $480 \times 480 \times 3$  的图像, 第二层为卷积层, 卷积核大小为  $7 \times 7$ , 步长为 2, 补零为 3, 数量为 64, 将输入下采样为原来一半。之后是批归一化层 (BN), 线性激活层 (ReLU), 以及步长为 2 的最大池化层。最大池化层之后是残差模块 (Residual Blocks), 在 ResNet50 中残差模块分为 4 组, 每组数量分别为 3、4、6、3。模块大小由  $1 \times 1$ 、 $3 \times 3$ 、 $1 \times 1$  的 3 个卷积核组成, 这样设计是为了在保证精度的同时降低参数量, 从而减少计算开销, 模块结构如图 3 所示。模型保留了除预训练 ResNet50 网络中末端平均池化层 (Average Pooling) 及全连接层 (Fully Connecting) 之外的所有层。这两层网络由解码器代替, 解码器由上采样 (Upsampling) 模块组成, Res-Unet 采用了 5 次上采样。每次执行上采样 (反卷积) 时, 特征图空间分辨率加倍通道数减半, 并最终增加到输入层原图大小, 这样整个网络就实现了端到端的输入。每个上采样模块由一个内核大小为  $2 \times 2$ , 步长为 2 的转置卷积层、一个拼接层 (Concat)、一个批归一化层 (BN) 和一个 ReLU 激活层组成。同时, 在上采样过程中, 将下采样块与其相应的具有相同空间分辨率的上采样块相连, 这样保证了最后恢复出来的特征图融合了更多的低层次的特征, 也使不同尺度的特征融合, 丰富了语义信息。之后进行两次卷积操作, 卷积核大小为  $3 \times 3$ , 步长为 1, 补零为 1, 共同作为下一上采样块的输入。在输出端, 使用  $1 \times 1$  卷积层来获得具有与类别数相同的输出通道的特征图。网络的具体实施细节如表 1 所示。

## 2 数据集与环境参数配置

### 2.1 建立数据集

本文针对混凝土裂缝、剥落、泛碱、孔洞、露筋五种常见病害, 建立的数据集包含 3 个来源:

- (1) 使用相机在不同光照条件下拍摄了 200 张分辨率为  $4608 \times 3456$  的房屋、桥梁相关病害图像;
- (2) 从互联网收集了 500 张分辨率不等的病害图像;
- (3) 选取部分 LI SY 等<sup>[10]</sup>开源的病害数据集。原始数据集如图 4 所示。

将滤波之后的图像进行人工像素级标注, 采用 PixelAnnotationTool 软件制作真值标签 (The Ground Truth)。如图 5 所示。为了生成更多数据

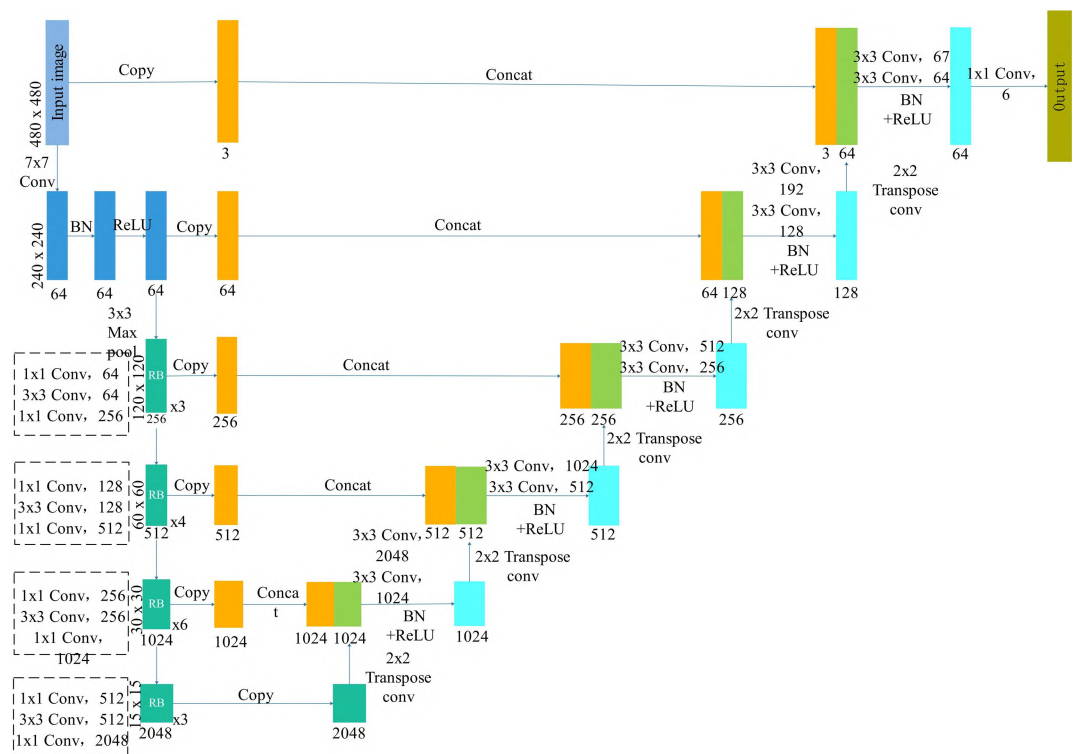


图 2 Res-Net 网络结构

Res-Net 的具体实施细节

表 1

层数	类型	卷积核大小	步长	填充	通道数	输出大小	通道数
1	Input	—	—	—	—	480×480	3
2	Conv_1+BN+ReLU	7×7	2	3	64	240×240	64
3	Maxpool_1	3×3	2	1	—	120×120	64
4	Residual blocks_1(RB)	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 3$	1	1	64 64 256	120×120	256
5	Residual blocks_2(RB)	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 4$	1 2 1	1	128 128 512	60×60	512
6	Residual blocks_3(RB)	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 6$	1 2 1	1	256 256 1024	30×30	1024
7	Residual blocks_4(RB)	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 3$	1 2 1	1	512 512 2048	15×15	2048
8	Transpose Conv_1	2×2	2	—	1024	30×30	1024
9	Concat_1	—	—	—	—	30×30	2048
10	Conv_block1	3×3	1	1	1024	30×30	1024
11	Transpose Conv_2	2×2	2	—	512	60×60	512
12	Concat_2	—	—	—	—	60×60	1024
13	Conv_block2	3×3	1	1	512	60×60	512
14	Transpose Conv_3	2×2	2	—	256	120×120	256
15	Concat_3	—	—	—	—	120×120	512
16	Conv_block3	3×3	1	1	256	120×120	256
17	Transpose Conv_4	2×2	2	—	128	240×240	128
18	Concat_4	—	—	—	—	240×240	192
19	Conv_block4	3×3	1	1	128	240×240	128
20	Transpose Conv_5	2×2	2	—	64	480×480	64
21	Concat_5	—	—	—	—	480×480	67
22	Conv_block5	3×3	1	1	64	480×480	64
23	Output_Conv	1×1	1	—	6	480×480	6

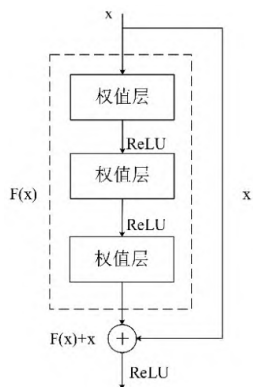


图3 残差模块



图4 原始病害数据集



图5 PixelAnnotationTool 软件标注真值标签

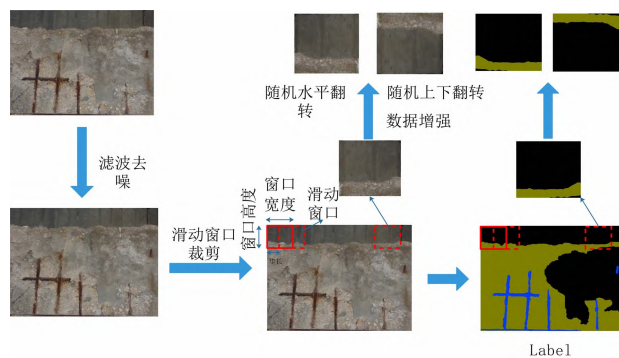


图6 图像数据预处理流程

集,利用滑动窗口技术对图像进行切割,窗口大小取  $1000 \times 750$ ,步长取 500,具体操作流程如图 6 所示。经过切割,一张  $4608 \times 3456$  的图像可以生成 114 张子图像,对于分辨率小于  $1000 \times 750$  的图像,采用线性插值方法调整图像大小为  $1000 \times 750$ 。经过切割调整之后,共产生 23500 张分辨率为  $1000 \times 750$  的图像,其中相当部分图像在切割时只包含背景,经过人工筛选,选择了 10000 张图像,随机进行水平翻转或上下翻转,最终得到 13200 张数据集用于训练。部分数据集及真值标签如图 7 所示。

将 13200 张病害图像数据分为训练集、验证集、测试集。其中 90% 用于训练和验证,10% 用于测试。具体的分配比例如表 2 所示。

## 2.2 参数配置

模型训练参数配置如表 3 所示。由于计算机配置限制,批次大小为 4,优化函数采用的是随机梯度下降算法 (SGD)、动量因子为 0.9,损失函数采用交叉熵损失 (Cross Entropy),初始学习率为 0.001,每 20 轮减小 10%,通过逐渐减小学习率来保证网络梯度下降时逐渐趋于最优解。

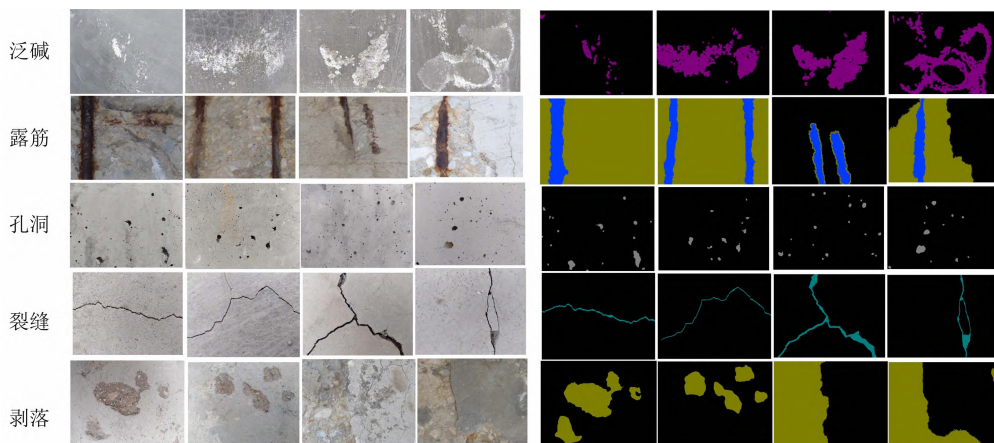


图7 部分病害数据集及标签

病害数据集分配明细 表 2

病害类型	训练集	验证集	测试集
裂缝	2190	283	278
剥落	2100	257	264
泛碱	2100	275	281
露筋	2110	261	256
孔洞	2010	268	267
合计	10510	1344	1346

超参数配置 表 3

超参数	值	超参数	值
Batch Size	4	Epochs	50
Learning rate	0.001	Loss	CrossEntropyLoss
Optimizer	SGD	Momentum	0.9
weight-decay	0.0001	—	—

### 2.3 评估标准

为了验证模型检测效果,采用 mIoU (平均交并比)、mPA (平均像素精度) 作为评价指标对比分析训练效果。

$$mPA = \frac{1}{k+1} \sum_{u=0}^k \frac{P_{uu}}{\sum_{v=0}^k P_{uv}} \quad (1)$$

$$mIoU = \frac{1}{k+1} \sum_{u=0}^k \frac{P_{uu}}{\sum_{v=0}^k P_{uv} + \sum_{v=0}^k P_{vu} - P_{uu}} \quad (2)$$

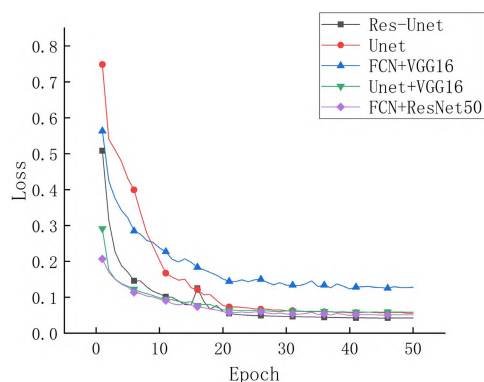
其中,平均像素精度 (mPA) 为计算每个类内被正确分类像素数的比例,再求所有类的平均;平均交并比 (mIoU) 为计算每个类真实值与预测值的交集和并集之比,之后求平均。 $k+1$  表示加入背景病害类别数,  $P_{uv}$ 、 $P_{vu}$  表示类别为  $u(v)$  的像素被预测为类别  $v(u)$  的数量,即假阳性 (False Positives, FP), 与假阴性 (False Negatives, FN),  $P_{uu}$  表示为像素预测正确的数量,即阳性 (True Positive, TP)。

### 3 模型训练与结果分析

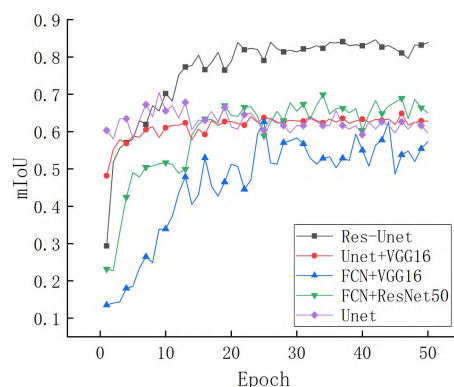
本文将 Unet 模型与 ResNet50 网络相结合,提出了一种改进的 Res-Unet 模型,为了验证模型的性能优势,将 Res-Unet 模型分别与原 Unet 模型、Unet+VGG16 模型、FCN+VGG16 模型和 FCN+ResNet50 模型对比。在相同超参数训练下,模型迭代 50 轮后区域收敛。损失函数、mIoU、mPA 曲线如图 8 所示。

由图所示,在训练前期随着迭代轮数的增加损失函数快速下降,这说明模型参数向着最优解

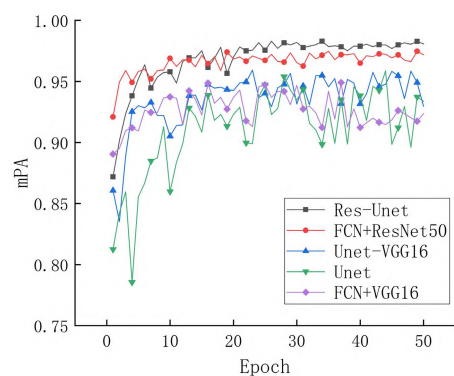
收敛,在大约 10 轮迭代之后,损失函数值降至 0~0.1 之间,30 轮之后 Res-Unet 模型损失函数逐渐收敛于 0.04 之间,在所有模型中最低。



(a) 损失函数曲线图



(b) mIoU 曲线图



(c) mPA 曲线变化图

图 8 损失函数、mIoU 及 mPA 曲线变化图

由 mIoU 及 mPA 曲线图可知,Res-Unet 模型有着良好的表现,在 mIoU 中,Res-Unet 模型平均交并比达到了 86.3%, 远高于第二位 FCN+ResNet50 模型 (67.4%), 在 mPA 曲线中,Res-Unet 模型平均精度达到了 98.5%, FCN+ResNet50 模型精度略低于 Res-Unet 模型,但也达到了 97.4%, 表现最差的是 Unet 与 FCN+VGG16, 分别为 91.7%、92.3%。模型训练

完成后,在测试集上分别验证了模型效果,测试结果如图 9 所示。

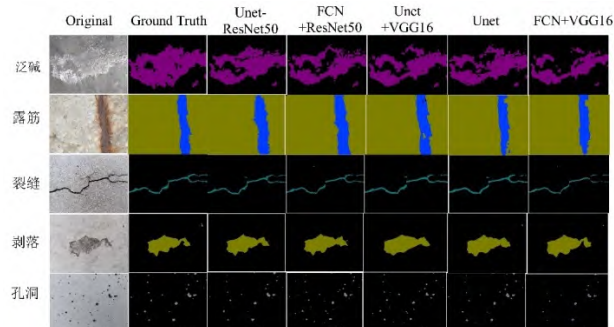


图 9 模型测试输出效果对比

#### 4 实际应用

将 Res-Unet 模型应用于实际场景检测中,以验证模型的鲁棒性。采用现场拍摄的原始数据集检测,其中所检测图像均未为未训练图像,由于相机拍摄图像分辨率过高,不能直接输入模型,故采用滑动窗口技术,在图像中依次以  $480 \times 480$  窗口扫描图像,一张分辨率  $4608 \times 3456$  的图像平均检测时间约为 3s,检测结果如图 10 所示。

由上述检测结果可知,对于裂缝,基本能够保证提取完整的裂缝骨架并检出细微裂缝;对于剥落和露筋现象在桥梁中居多,往往同时出现,虽都能检出,但当背景复杂或剥落与正常混凝土背景相似时检测效果稍差;对于泛碱与孔洞,与背景混凝土反差较明显,能较好检测出病害轮廓位置,部分会存在漏检现象。综上,Res-Unet 网络能够检测出所有的病害位置,整体效果良好,但算法在复杂背景中的检测尚有待改进,后续研究会继续补充样本集,增加工况场景,提升模型性能。

#### 5 结语

本文通过建立一个混凝土常见病害(裂缝、剥落、泛碱、露筋、孔洞)数据集,提出一种改进的 Unet 网络,通过调整超参数来训练模型,最终实现了图像数据像素级分割,得到以下结论:

(1) 通过人工采集混凝土病害图像数据,并对数据进行人工标注,预处理。构建了可用于训练的混凝土多病害数据集。

(2) 利用 ResNet50 网络进行特征采样,基于 Unet 框架提出 Res-Unet 网络模型,经过调参优化,在 50 轮迭代训练之后,模型趋于收敛,mIoU 能够达到 86.3%,mPA 能够达到 98.5%。均优于原 Unet、FCN 模型与 VGG16 作为编码器的模型。

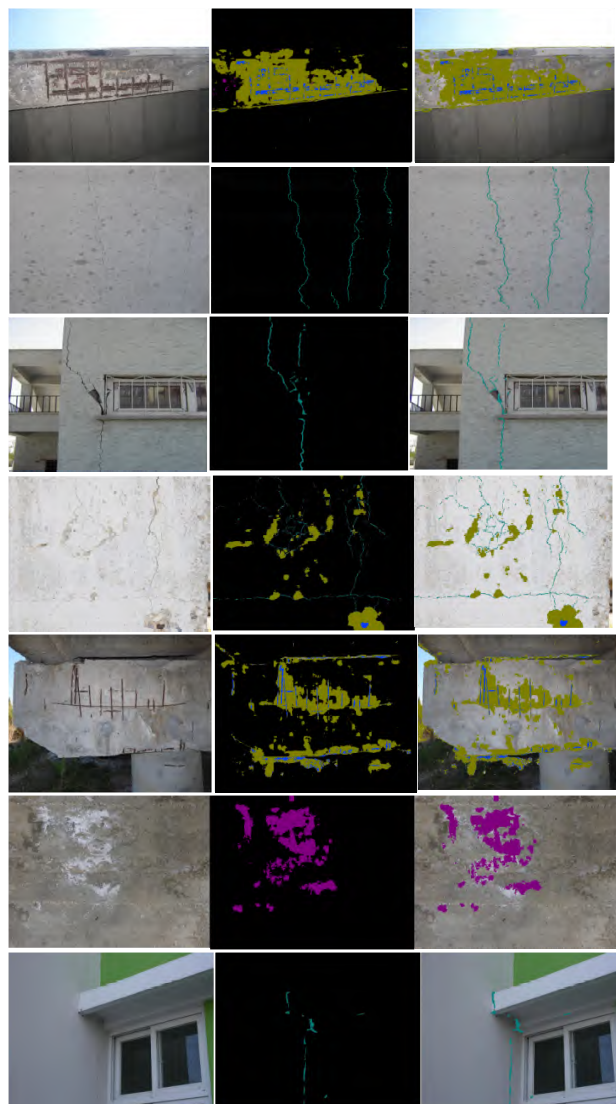


图 10 实际场景检测效果

(3) 为保障混凝土表面病害检测人员安全、减少工作量实现高效自动化检测以及混凝土结构的维护、保养提供了技术支持与理论指导。

#### 参 考 文 献

- [1] 周颖,刘彤. 基于计算机视觉的混凝土裂缝识别[J]. 同济大学学报(自然科学版), 2019, 47(9): 1277-1285.
- [2] 王丽. 混凝土坝裂缝识别的数字图像处理算法研究[D]. 大连: 大连理工大学, 2019.
- [3] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [J]. The IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBoxDetector [C]//14<sup>th</sup> European Conference on Computer Vision. Amsterdam: University van Amsterdam, 2016: 21-37.
- [5] RONNEBERGER O, FISCHER P, BROX T. U-Net:

- Convolutional networks for biomedical image Segmentation [C]//International Conference Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer Cham, 2015: 234-241.
- [6] 韩晓健, 赵志成. 基于计算机视觉技术的结构表面裂缝检测方法研究[J]. 建筑结构学报, 2018, 39(s1): 418-427.
- [7] CHA Y J, CHOI W, BUYUKOZTURK O. Deep learning-based crack damage detection using convolutional neural networks [J]. Computer-Aided Civil and Infrastructure Engineering. 2017. 32(5): 361-378.
- [8] PARK SOMIN, et al. Patch-based crack detection in black box images using convolutional neural networks[J]. 2019, 33(3): 04019017.
- [9] 李国耀, 王腾. 基于形态学处理与特征分析的混凝土裂缝检测研究[J]. 建筑结构, 2020, 50(S2): 529-533.
- [10] LI S Y, ZHAO X F. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network [J]. Computer-Aided Civil and Infrastructure Engineering. 2019, 34(7): 616-634.